Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 96 (2013) 2320 – 2331

Procedia Social and Behavioral Sciences

13th COTA International Conference of Transportation Professionals (CICTP 2013)

Research on Freight Truck Operation Characteristics Based on GPS Data

Fu Yanhong[a], Shi Xiaofa[b,*]

[a] *Graduate Student of School of Transportation Engineering, Tongji University, 4800 Cao'an Road, Shanghai, P.R.China; PH (86).021-69583775 ;Email:twinkle6fu@163.com*

[b] Assistant Professor of *School of Transportation Engineering, Tongji University, 4800 Cao'an Road, Shanghai, P.R.China; PH (86).021-69583775 ;Email:*xfshi999@163.com

Abstract

The GPS data of trucks, which records the truck operation state in real time, can provide a large number of basic information for freight transportation planning and management. The paper studies the Application of GPS data mining in freight transportation. The paper analyses the freight truck utilization, the spatial-temporal features of truck trips, the flux, flow of freight corridors and some related traffic information by matching the truck tracks into Google Earth, coordinate transformation, trip division and other methods. Through comparing the results of sample data analysis with the actual situation, the method to analysis freight truck operation characteristics based on GPS data is proven with strong reliability.

Key words: GPS data; spatial-temporal feature; trip division

1. Introduction

Freight transportation is essential for urban economic and social activities, the basic requirement for survival in cities. Freight traffic and passenger traffic are both the major components of urban traffic. However, urban planners and traffic engineers focus mainly on the passenger transportation and the needs of inhabitant trips, taking into account only cars and buses, rarely involving freight cars. With the high-speed economic development, freight traffic is increasing rapidly, so experts begin to notice that the cause of traffic congestion isn't only passenger vehicle, but the increasing number of trucks compete the limited road space with cars and buses[1]. Therefore, better planning and managing freight transport is one of the key measures to improve urban traffic. Further, freight traffic survey is an important basic work of freight transportation planning and management, which directly affects the reliability of the program.

Freight traffic surveys primarily include freight vehicle trip utilization, trip OD distribution, spatial and temporal characteristics, the main corridors and other truck operation information. The traditional freight traffic surveys are typical enterprise questionnaire survey to obtain the freight transport demand and situation of

* Corresponding author. Tel.: 86-13816331510,86-21-69583699; fax: 86-21-69583775.
E-mail address: xfshi999@163.com.

doi:10.1016/j.sbspro.2013.08.261

enterprise, and the traffic volume and speed survey in specific road sections. The methods are difficult to implement widely, and the reliability of data is low, because it costs a lot of human and financial resource, while involving commercial confidential information. The paper discusses the method to analysis truck utilization, trip characteristics, flux and flow of freight channels and other freight traffic characteristics by mining the floating truck data.

In recent years, researches on freight survey pay more attention in the freight volume forecast, that the qualitative and quantitative forecasting methods are more than 300 species at home and abroad. The methods concentrated in time series, regression analysis, neural networks, grey model and other models[2]. Xijun Ge studied spatial distribution, characteristics and development of logistics demand in cities agglomeration region with the node city as the basic unit and integrating urban social, economic and geographical and other relate factors[3]. In considering the concentrate degree of industrial distribution, Liyuan Liu suggested to selecting typical enterprises to investigate on characteristics of freight according to the different industries and the scales of production, and the freight distribution would be derived by sampling expansion combining survey data with macro-statistics[4]. These studies focus primarily on distribution and forecast of freight transport demand, lack of study on characteristics and the effect of freight traffic, so it's difficult to directly support the freight transport planning.

GPS floating car technology increasingly apply in transport with the development of vehicle position and wireless communications technology, which mainly concentrated in the prediction of road travel time, vehicle dynamic route guidance, publishing the real-time road network traffic information and detecting traffic incident in network, as well as analysis of vehicle operation characteristics combined with taxis and buses[5]. Although freight vehicle is also a kind of floating car, the GPS data mainly use to determine the location of the goods, while the research on analysis of freight vehicle running characteristics combining the data is few. The paper explores the method to mining trucks running characteristics to support freight transport planning and management by making use of the GPS data of trucks.

2. The Theory and Methods of GPS Data analysis

GPS Equipped Probe vehicle is currently an advanced technology for collecting road traffic information in ITS field at the international level. The vehicle information (car ID, time, coordinates, instantaneous speed, direction and other information) can be transferred to information center in real time by installing GPS device and wireless communication equipment in a car. The traffic information of vehicle operation characteristics can be extracted after extensive vehicle information collection and process [6].

2.1 Structure of GPS Data

The data collected by GPS includes car ID, time, place coordinates, instantaneous speed, direction, running status and other information. The data structure used in this paper is shown as table 1.

Table 1. The GPS data structure of floating car

Field name	Data type	Sample	Illustration
ID	Char(9)	粤 B012GSD	The uniquely identify of car
Longitude	Numeric(17,15)	114.21717071533203	Degree as unit, retain 15 decimal places
Latitude	Numeric(17,16)	22.625667572021484	Degree as unit, retain 15 decimal places
Time	Nvarchar(18)	2011-10-01 15:52:06	15:52:06, 10/1/2011
Time difference	Bigint(13)	1319481044163	Milliseconds from the base time 0:00:00, 1/1/1970
Instantaneous speed	Int(3)	45	45km/h
Running status	boolean	0,1	0 is empty, 1 is heavy

2.2. Data filtering

Due to system internal or external factors, there are some unreasonable data in the GPS raw data, which are main large errors or incorrect data. The mistakes shows as :①'The drift of data', which is GPS track points drift, it is serious when trucks driving in elevated or tunnel.②'The illusion of data', which is the truck stop actually, but the data returned from the system are not the same. It appears the illusion of truck moving because the return information is fluctuation within more than 10 meters or dozens of meters. ③'The jump of data', which is there are big gap between adjacent data when trucks drive through communication base station or large electric power equipment, that cause a significant impact on truck location and speed[7]. In addition, there are some mistakes of return time and vehicle status occurred when GPRS transmit the information to centre.

Therefore, it is necessary to preprocess the data to enhance the credibility of the data, reduce the uncertainty and errors. The preprocessing procedure in the paper is: firstly, delete the wrong time data; secondly, delete the duplicate data; finally, delete the data which return status is not consistent with the actual state.

2.3. Map matching

Map matching is a positioning method based on software technology, whose basic idea is to combine the vehicle track from the trajectory-positioning device with electronic map, and determine the location of vehicle. In the paper, because the data is come from the running trucks through the port, so the travel distances are relatively long which don't limit in a city, even cross the province, which takes much effort and time to make such a large electronic map. Therefore, the paper discusses the possibilities of using Google Earth as map matching database.

Google Earth is earth software of three-dimensional visualization introduced by Google, which has a huge database and data source with the terrain, buildings, environment, traffic and other information anywhere around earth, and displaying in different graphical hierarchical. It plays more powerful features than information browsing through adding a variety of information into Google Earth by making different data layers[8].

Nowadays, Google Earth only supports KML format files based on XML grammar and file, and KML stands for Keyhole Markup Language. KML is used to describe, preserve, exchange of geography information and map objects, such as points, lines, images, and display in Google Earth. The paper matches the GPS information into Google Earth by transform the ID, geographic and time of trucks into KML files.

2.4. Coordinate transformation

The coordinate systems in GPS and Google Earth are all the WGS-84 coordinate system (World Geodetic System-84), which is a geocentric coordinate, so it is hard to calculate the distance between the origin and the destination of trip from longitude and latitude. It is necessary to transform the coordinate, which is to cast the longitude and latitude coordinates in WGS-84 as a benchmark from GPS receiver to plane coordinates. The specific steps take 7-element Helmert to convert the latitude and longitude data which are reflect the place of the running truck to x and y coordinates in the plane coordinates[9]. Table 2 shows the results of transforming.

Table 2. The results of coordinate transformation

Longitude	Latitude	X_jd	Y_wd
114.54329681396484	22.862428665161133	165753.544604105	54345.9302649274
114.54329681396484	22.862428665161133	165753.544604105	54345.9302649274
113.95649719238281	22.7362003326416	105139.870378113	40803.1293220334
113.68820190429688	22.844600677490234	78129.2681550535	52438.7462551482
113.80580139160156	22.794599533081055	89918.3160341557	47148.5712892003

3. Freight Trucks Operation Characteristics

The example data of this paper is GPS data of 4439 trucks through Yantian Port which logged trucks running state all the day during October 24th to 30th, 2011. Figure 2 shows the data processing flow of the paper.

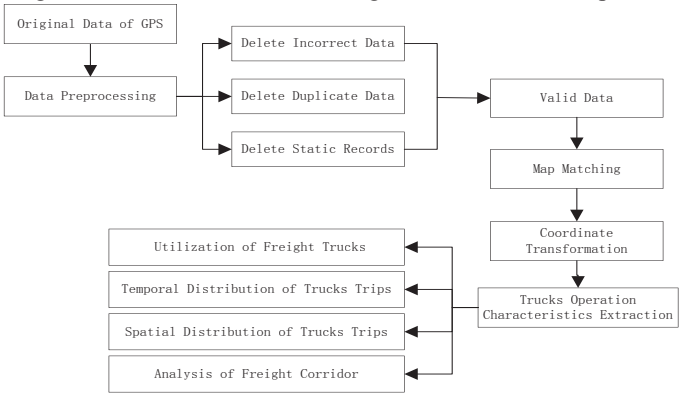

Fig.1. The data processing flow

After processing, the original data records reduced from 108.2 million to 60.3 million that loss rate of the data is 44.3%. The quality of valid data is to examine whether the interval of records can support to extract the freight car operation characteristics. Figure 2 is analysis result of valid records time interval.

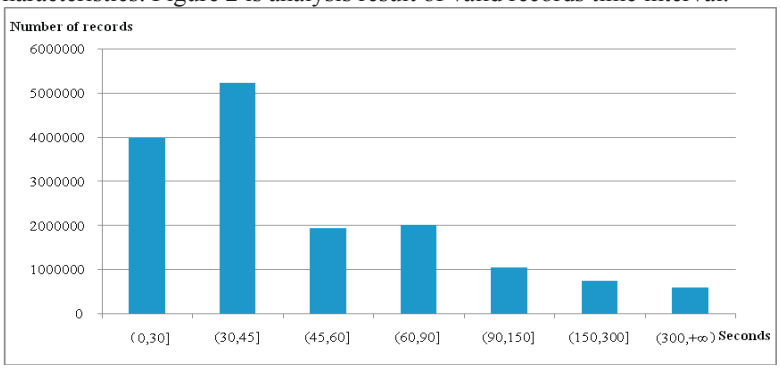

Fig.2. The frequency of the valid records time interval

The return record time interval of trucks is concentrated in less than 45 seconds, which reaching 59.38%, and the number of records whose time intervals within 90 seconds reaches 84.84%, so the number of long time interval records is very small. Assume that the speed of the truck is 60km/h and the distance is 750 meters in 45 seconds and 1500 meters in 90 seconds. Taking into account the roads of trucks travelling are mainly intercity expressway, urban expressway and main road, whose intersections and entrances are few, so the result of the actual track of trucks, which is obtained by the distance, is with high precision and credibility.

3.1. Analysis of Freight Trucks Utilization

The license plate number of trucks is the identification to distinguish a car different from other trucks, so by doing statistics the number of unduplicated license plate number every day, we can get the number of operation trucks on that day.

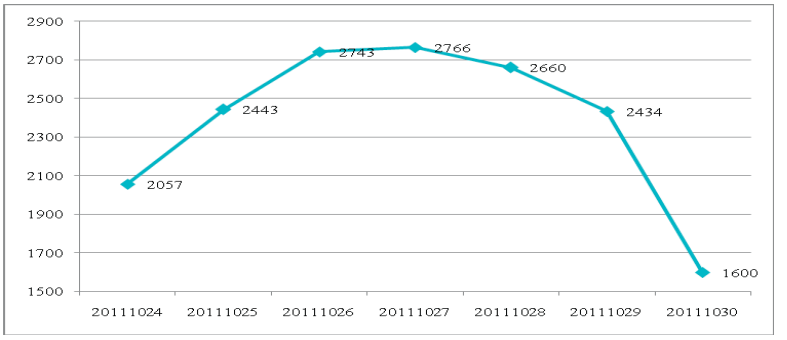

Fig.3. The number of Freight trucks through Yantian port

According to statistics, the number of trucks in and out of Yantian port is around 2000 every day accounted for 53.8% in total number of survey trucks, and this means that the average number of running trucks is over half per day. The sample data has a certain level of representativeness for research on port cargo flow and trucks running attribute.

Freight truck utilization is the ratio of operation trucks per hour in survey trucks per day. The result shows that the utilization of trucks is 42.5% per day, 45.5% during working day and 35.2% on the weekend. This means that the using strength of trucks has similar characteristics, and the usage is more efficient in working days.

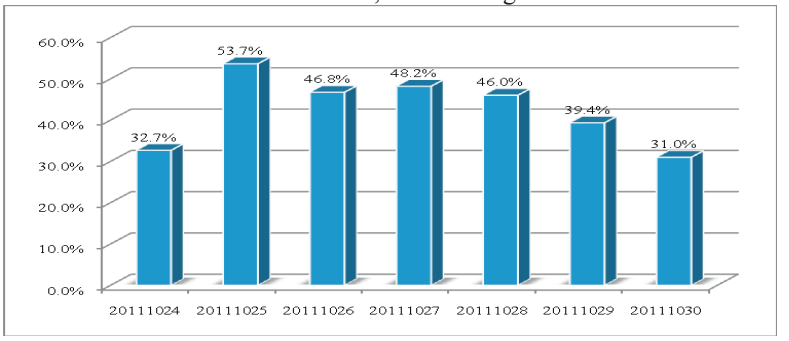

Fig.4. The utilization of survey trucks average day

3.2. Temporal distribution of freight trucks trip

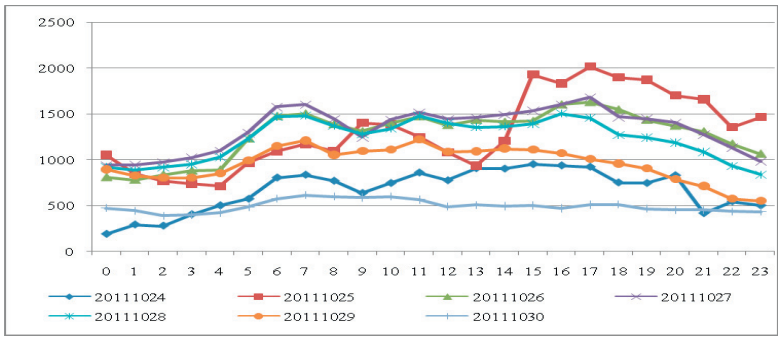

Fig.5. The temporal distribution of operation trucks

The query result shows that the time of trucks running has similar characteristics on the same period of different days. Trucks travel time is stable throughout of day, and the peaks and valleys of trip are not obvious. Trucks travel time is concentrated in 6:00-21:00, which appears a small peak in 6:00-8:00. The freight traffic volume reduces in 8:00-9:00, and is low in 0:00-5:00 and 21:00-23:00. According to 2011 Shenzhen annual report on transport development, since the work schedule changed to from 9:00 am to 6:00 pm, the morning peak of trip is 7:00-9:00 and the evening peak is 17:00-19:00 in the whole road network, but the morning peak is 7:45-8:45 in downtown area. Obviously, the morning peak of freight transport is earlier than commuter traffic, which is to avoid commuter traffic reducing effect.

3.3. Spatial distribution of freight trucks trip

(1) Trip division of freight trucks

Due to the GPS data record uninterruptedly running status of freight vehicles, these records may be from the same trip or different trips, so it is necessary to divide trip. The paper selects based on time-divide method to split trip. Based on time-divided method is to divide trip in accordance with the time distribution law of records, whose basic idea is that the time of records has centralization in the same trip and variation between different trips for a car. When the discreteness of time reaches a certain level, it is considered the record from next trip[10]. The important point of this method is to determine the time threshold. Because the time to wait for loading and unloading goods is long, the time threshold to splitting trip for freight trucks should not be too short. Table 3 provides a result of trip division from different trip time thresholds.

Table 3. The average trip time of freight trucks from different time thresholds

Date	Average trip time when TS=60min	Average trip time when TS=90min	Average trip time when TS=120min
20111024	2.346	1.779	1.416
20111025	1.948	1.485	1.155
20111026	2.336	1.665	1.314
20111027	2.187	1.661	1.334
20111028	2.558	2.042	1.695
20111029	1.798	1.380	1.101
20111030	1.476	1.133	0.932
Average trip time	2.093	1.592	1.278
Average trip time on working day	2.275	1.726	1.383

Remark: TS in Table.3 means time thresholds.

It is considered the first trip that the truck went to Yantian port delivery and the second trip left from port, so the average trip time is more than two. According to table 3, it is more reasonable to split trip by 60 minutes as

time threshold. Specific process of trip division shows as Figure 6. All records in the same trip appears after splitting trips, and the position calculated by latitude and longitude in the home and the end record of a trip is origin (O) and destination (D) of this trip.

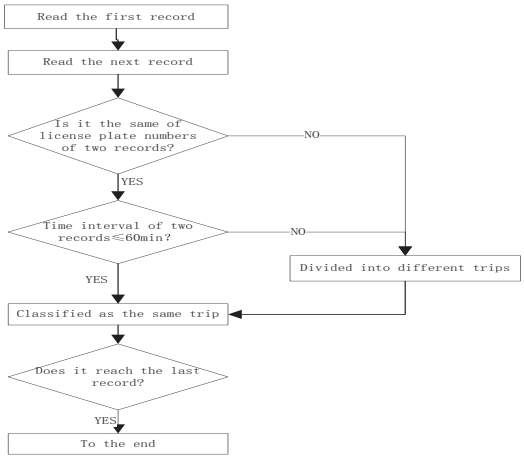

Fig. 6. Algorithm process of trip division

(2) The average linear distance of truck trips

The latitude and longitude of trip OD can convert to coordinates of XOY by the program of coordinate transformation, and then the linear distance of trips would be calculated by distance formula between two point. The result shows as Figure 7.

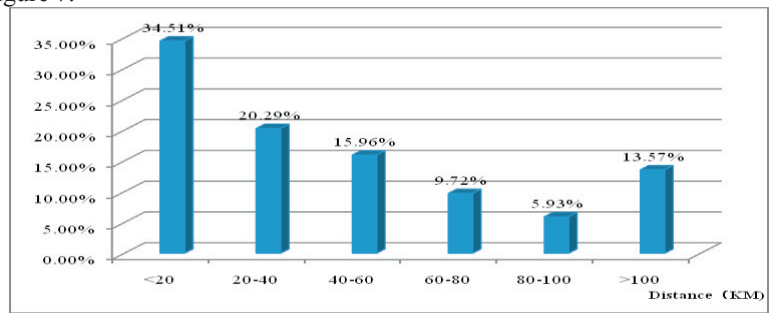

Fig. 7. The average linear distance of truck trips

According to Fig.8, the trip times in distance less than 40km, 60km, 80km reaches respectively 54.8%, 70.8% and 80.48% in all. A point of OD is in Yantian Port therefore the distance represents the distance between the origin or destination of freight and Yantian Port. Figure 8 is concentric circles which the center is Yantian Port, so the origin or destination of truck running are main in Shenzhen, Dongguan, Huizhou, Guangzhou, Zhongshan, Zhuhai and Foshan.

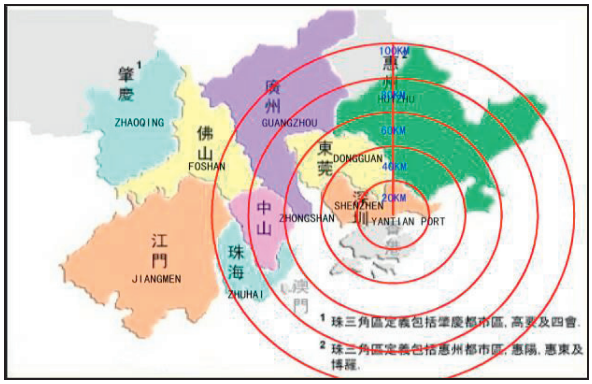

Fig. 8. The distance circle of Yantian port

In table 4, the average trip distance of survey trucks is 55.7km, that is say the travel range of investigated trucks is mainly in Shenzhen, Dongguan and Huizhou.

Table. 4. The average trip distance of freight trucks per day

Date	The average trip distance (KM)	Date	The average trip distance (KM)
20111024	45.89451	20111028	62.80913
20111025	39.4438	20111029	56.98729
20111026	64.73512	20111030	55.49625
20111027	64.59683	The average trip distance	55.70899

(3) The OD distribution of freight truck trips

According to the record filter method, one end of truck trip is Yantian port and the other end is origin or destination of trip. The method to getting OD distribution of freight vehicles is to match the latitude and longitude of the OD and traffic zones, which are the administrative areas of cities in Google Earth, but the traffic zone of Shenzhen doesn't include Yantian port. The result of map matching shows in Figure 9.

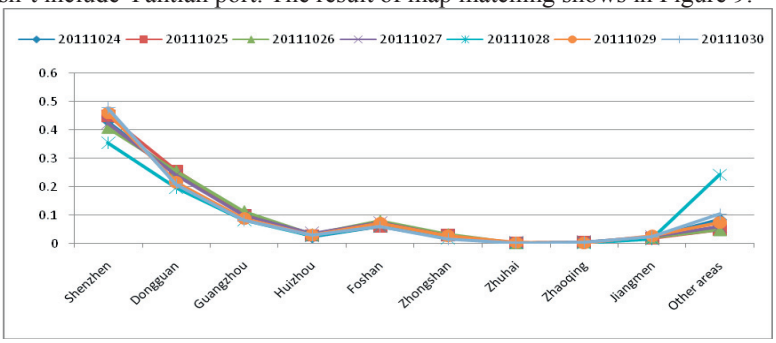

Fig. 9. The trip distribution of freight trucks through Yantian port

The travel scope of survey freight trucks through Yantian port is Shenzhen, Dongguan, Guangzhou, Foshan, Huizhou and Zhongshan, in which the total number of trucks in these six cities accounts for 87.7 percent. According to public documents, the largest source of freight in Shenzhen port came from Pearl River Delta area, especially Shenzhen, Huizhou, Dongguan, Zhongshan, Foshan, that the foreign trade volume of this five cities accounts for 75.2% in Guangdong Province[11]. By contrast, the trip OD distribution acquired from GPS data agree with the actual situation, so the method to extract OD distribution of freight truck has a strong reliability.

(4) The Freight Corridor of Yantian Port

① The main freight corridors of port

The planning corridors for goods collecting and distributing in Yantian port are Yanpai Expressway, Yanhui Highway in the north of port, Yanba Expressway in the northeast of port, and Luosha Road in the southwest of port, as shown in Figure 10.

Fig.10. The planning corridors for goods collecting and distributing in Yantian port

Figure 11 shows the map matching result in Google Earth by transforming the excel files which save the GPS track points of trucks to the KML files. We can see that Yanpai Expressway, Yanhui Highway and Yanba Expressway are the main channels in and out of the Yantian Port, but the trucks are scarity in Luosha Road, because Luosha Road is closed to freight traffic since 2006.

Fig.11. Freight trucks GPS track for entering into and leaving port

②The freight trucks flux and flow direction in different channels

Traffic organization for collection & distribution traffic and transit traffic of Yantian Port is " enter from the east and leave from the east, enter from the north and leave from the north", so Yanpai Expressway is primarily

responsible for trucks from Shenzhen, Dongguan, Guangzhou and the west areas of the port, while Yanhui Highway and Yanba Expressway is responsible for trucks from the east of Shenzhen, Huizhou and Shantou and so on.

Fig.12. Freight traffic organization for collection & distribution in Shenzhen Port

By statistics trip OD of trucks and the channels for trucks moving in and out of Yantian Port, we can analysis main trucks distribution in each channel.

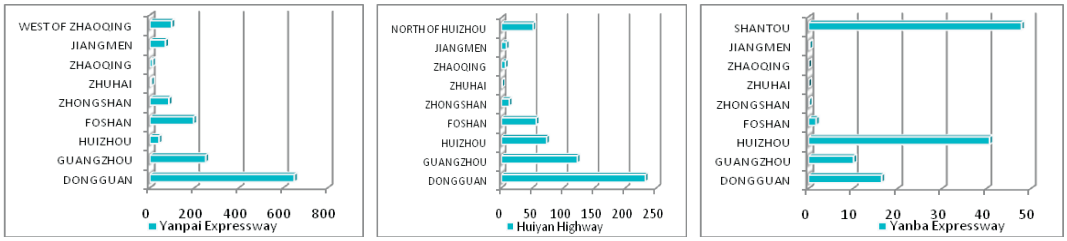

Fig.13. The spatial distribution of trucks in each channel

As shown in figure 13, trucks through Yanpai Expressway are from the west of Shenzhen including Dongguan, Guangzhou, Foshan, Zhongshan, Jiangmen, etc; The trucks through Huiyan Highway are from Dongguan, Guangzhou, Huizhou and other cities; The trucks through Yanba Expressway are from the east of Shenzhen, such as Huizhou, Shantou. The flux and flow direction of each channel which come from GPS data agree with the actual measure that traffic organization in Yantian Port.

③ The temporal distribution of freight traffic flow in each channel

The temporal distributions of traffic flow in each channel can analysis the peak period of the channels, which is the important indicator of road service level.

According to the truck flow statistics of each channel every day, there is less fluctuations and no obvious peak in trucks flow of each channel. The small peak in Yanpai Expressway and Huiyan Highway appears in 5:00-7:00 am, and the flow significantly reduces in 7:00-9:00 am. This is because that trucks entering into or leaving Yantian Port through Yanpai Expressway or Huiyan Highway must move through the centre of Shenzhen, so the truck choose to avoid the urban commuter traffic during the peak. Yanba Expressway belongs to crossing channel, so the truck flow increases after 9:00.

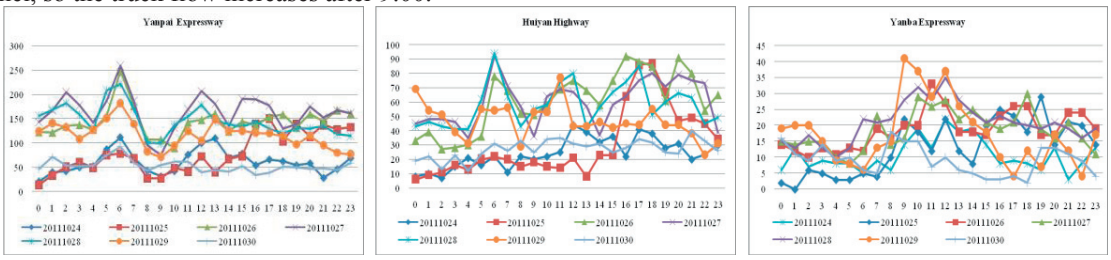

Fig.14. The temporal distributions of trucks in each freight channel of Yantian Port

4. Conclusion

The paper discusses the method of analysis of freight truck operation characteristics in GPS data by data filtering, map matching in Google Earth and coordinate transforming. It is proven that the method is practicable, and the reliability of the traffic information is strong which can provide reference information for freight transport planning, combining with the example data of truck running through Yantian Port. The main conclusions are as follows:

(1) The GPS data is suitable for analysis of freight truck operation characteristics. The paper analyses freight truck utilization, trip OD distribution, flux and flow of freight corridors and other characteristics by the GPS data to lay the foundation for further study on freight traffic. The method makes up for the disadvantages of time-consuming, laborious and expensive in traditional survey, with strong enforcement.

(2) It is credible and easy operational to match map-making use of Google Earth. The article researches the truck track, trip distribution, and freight corridors by the powerful database — Google Earth, which can save much time to draw an electronic map and avoid the errors during map matching. The advantage of this method is less complicate and can be widely used.

(3) The result of trip division in freight truck records is reasonable. It is essential to split the trips when analyzing the truck running characteristics. The method of dividing trips is to discuss the way of determining the time threshold based on time agglomeration theory. The result is in line with the actual situation, so the method has reference.

The GPS data of truck contains a lot of traffic information during trucks running, which not only monitoring the vehicle in real-time, but also providing much traffic information for freight traffic planning and management. Of course, the method has requirements for the quality and quantity of the collected data, which is to be improved

by the development and breakthrough of the related technologies to GPS. At the same time, the methods of data filtering, standardization and accuracy all need to be further studied and optimized.

Reference

[1] Xiang He. Temporal-spatial Correlation Analysis of Urban Road in Information Environment[D]. Shanghai: Tongji University, *2009.3*(5~10)

[2] Linlin Liang. Congestion Correlation Analysis of Urban Expressway Based on Data Mining[D]. Shanghai: Tongji University, *2012.3*(33~50)

[3] M Jain, B Coifman (2005), Improved Speed Estimates from Freeway Traffic Detectors, *Journal of Transportation Engineering*

[4] JIANG Gui-yan, GANG Long-hui, ZHANG Xiao-dong, WANG Jiang-feng (2004). Malfunction identifying and modifying of dynamic traffic data. *Journal of Traffic and Transportation Engineering*, 1. 3 - 5

[5] Chen.M., J. Xia, R. Liu. Developing a Strategy for Imputing Missing Volume Data, *Proceedings of the 85th Annual Conference of Transportation Research Board*, *2006.1*.

[6] Ya Sun. Study on Fixed Traffic Information Collection System Based on Data Quality [D]. Shanghai: Tongji University, *2008.7*(84~95)

[7] Ji Yushan, Wu Yongmin. On the Cointegration-model of the Relations between Industrial Structure and Economic Growth in China. *Contemporary Economic Research. 2006.6*

[8] Ya-qun HE, Guo-hong LAO, Chris E OSUCH, Wei-ran ZUO, Bao-feng WEN. Co-integration-based analysis of energy assurance for steady economic growth in China. J*ournal of China University of Mining and Technology. 2008.6* (pp. 250-254)

[9] Li Weifeng, Duan Zhengyu. Error Correction Model of Floating Car Data Based on Co-integration Theory.

[10] Tiemei Gao (2009). *Econometric Analysis and Modeling* (2th ed). Beijing: Tsinghua University Press (pp. 164-182)

[11] William H. Greene (2001).*Econometric Analysis* (4th ed). New York: Prentice Hall (pp. 748 - 760)